

## Research Article

# The Dorsal Differentiation of Velar From Alveolar Stops in Typically Developing Children and Children With Persistent Velar Fronting

Joanne Cleland<sup>a</sup>  and James M. Scobbie<sup>b</sup>

**Purpose:** This study has two key aims: first, to provide developmental articulatory norms for the alveolar–velar distinction in 30 English-speaking typically developing (TD) children; second, to illustrate the utility of the reported measures for classifying and quantifying the speech of children with a history of persistent velar fronting as they develop the contrast longitudinally.

**Method:** This study involved secondary data analysis of the UltraSuite corpus comprising ultrasound tongue imaging recordings of speech materials from 30 typical children and longitudinal data from five children with persistent velar fronting undergoing ultrasound visual biofeedback intervention. We present two new measures of coronal dorsal differentiation: KTmax and KT crescent area. These measures distinguish /k/ and /t/ by quantifying the magnitude of this distinction in absolute spatial terms (mm of linear dorsal difference). For the typical children, we report these measures in corner

vowel contexts. We then compare these to dorsal productions by the children with speech disorders, before, during, and after intervention.

**Results:** Both measures reliably distinguished /k/ and /t/ in TD children. There was an effect of vowel, with larger KTmax and KT crescent area in /a/ and /o/ vowel contexts than in an /i/ context. The children with persistent velar fronting showed KTmax values near zero before intervention, showing a complete merger between /k/ and /t/. During intervention, they showed variable KTmax values. Post intervention, they showed values within the range of typical children.

**Conclusions:** This study provides articulatory norms derived from ultrasound tongue imaging for the dorsal differentiation in alveolar and velar stops in TD children. By applying these norms to children with persistent velar fronting as they acquire this contrast, we see that /k/ is acquired in an articulatorily gradient manner.

A critical milestone in the acquisition of intelligible speech occurs when children become able to accurately produce all the consonants required for their language. This milestone is reached in a gradual way, with children often simplifying a consonant's production by omitting it or substituting another in its place (McLeod & Baker, 2017). One common substitution in young typically developing (TD) children is velar fronting (Dodd et al.,

2002). In this process, velar consonants are substituted by alveolar consonants, for example, /kat/ being realized as [tat] in Scottish English, leading to homophony and breakdowns in communication. Children's velar fronting is normally transient. A recent review of consonant acquisition across languages (McLeod & Crowe, 2018) shows that velars are acquired relatively early in a variety of languages: between ages 1;10 and 2;11 [years;months] for most children (75%–85% criteria). For English-speaking children specifically, velars are early acquired consonants, appearing between ages 2;0 and 3;11 for 90%–100% of TD children. Despite this, both preschool and school-aged children who have not yet acquired velars commonly present in speech-language pathology clinics. Velar fronting is a rule-based pattern and hence most often conceptualized in the literature as a phonological process. In fact, absence of velars in the phonetic inventory at 3 years of age is predictive of phonological disorder (Stoel-Gammon, 1996), and, therefore, velar fronting is normally

<sup>a</sup>School of Psychological Sciences and Health, University of Strathclyde, Glasgow, United Kingdom

<sup>b</sup>Clinical Audiology, Speech and Language Research Centre, Queen Margaret University, Musselburgh, United Kingdom

Correspondence to Joanne Cleland: [joanne.cleland@strath.ac.uk](mailto:joanne.cleland@strath.ac.uk)

Editor-in-Chief: Erika S. Levy

Editor: Cara E. Stepp

Received June 29, 2020

Revision received September 10, 2020

Accepted October 16, 2020

[https://doi.org/10.1044/2020\\_JSLHR-20-00373](https://doi.org/10.1044/2020_JSLHR-20-00373)

**Publisher Note:** This article is part of the Special Issue: Select Papers From the 2020 Conference on Motor Speech.

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

successfully treated with phonological interventions (McLeod & Baker, 2017).

However, it is not universally accepted that velar fronting can be ultimately explained through an abstract symbolic process such as the substitution of one category in the phonology by another. Ultimately, it may be articulatory, or motor constraint, explanations for this particular substitution that can best explain its etiology, and perhaps inform paths of remediation. In an influential paper, Gibbon (1999) describes “undifferentiated lingual gestures” (ULGs) in the speech of 12 children with “functional articulation disorders” (now known as speech sound disorders [SSDs] of unknown origin). These undifferentiated gestures occur in children who have been unable to learn how to independently control the coronal and dorsal regions of their tongue. Evidence for ULGs comes from mainly electropalatography (EPG) studies. The phenomenon is detected when a large amount of tongue–palate contact occurs during attempts at velar and/or alveolar stops. Contact may extend all the way from the alveolar region to the velar region. Relatively few recent studies of disordered speech have used EPG, while a different technique, ultrasound tongue imaging (UTI), has come into more common use (Sugden et al., 2019). ULGs can also be visualized using UTI if the probe is fixed, relative to the speaker’s head, and a hard palate tracing is superimposed. Doing so allows visualization in the midsagittal plane of increased contact between the tongue and the hard palate (Cleland et al., 2017). The greater availability of UTI to researchers means that it may now be possible to discover how common these ULGs are in various clinical populations, and indeed whether and how often they are seen in younger TD children who display velar fronting. Currently, such insights are not possible because previous instrumental studies have typically included only small numbers of children. Moreover, they typically have addressed the articulation of a limited clinical profile: more complex children for whom traditional phonological intervention has failed to remediate velar fronting (Cleland et al., 2015).

It is also, almost by definition, difficult to identify ULGs without articulatory data because, although ULGs of attempts at velar consonants are not alveolar substitutions per se, they may be readily heard as such, and hence transcribed as alveolars by speech-language pathologists. When a child consistently realizes /k/ in this way, but /t/ is produced as an alveolar stop, this is also known as a covert contrast. Covert contrast is a term used to describe the phenomenon where, despite a perceptual neutralization, there is any measurable acoustic or articulatory difference between attempts at different phonemes. This type of contrast is taken as evidence that the child has phonological knowledge of the contrast, but is realizing it in an unusual way that is not perceptible to the listener (Scobbie et al., 2000). In the case of a ULG, it is likely that this is because of a motor constraint. Articulatory work has also revealed that it is not just a tendency to use transcriptions that are impressionistically suitable for one or other of the potential phonological categories that is responsible for masking the true nature of these stops from the listener. It

could also be the case that the spatially undifferentiated gesture is also dynamically complex. Tongue–palate contact in a ULG is unlikely to be released completely, all at once. If contact is released first by the back of the tongue followed by a final release of contact at the front, on the alveolar ridge, this is likely to result in a [t] percept. Termed *articulatory drift* (Gibbon & Wood, 2002), this type of error is potentially common in children with SSDs and would account for a motor-based constraint resulting in a seemingly phonologically patterned error. Alternatively, other children with ULGs may display articulatory drift that is not always manifested in the same direction, resulting in the listener transcribing a variety of velar, alveolar, or palatal stops or coming to the conclusion that the child’s speech is highly variable.

In a more recent paper, McAllister Byun (2012) takes further the explanation that velar fronting is a phonological process brought about by specific constraints in children’s speech-motor control. McAllister Byun is interested in cases where the child does not front velars in all word positions, but instead shows positional velar fronting, with velars in onset or stressed positions more vulnerable to fronting. In this case, the child clearly does have the ability to differentiate the coronal and dorsal parts of the tongue some of the time: It is therefore tempting to conclude that this must be a truly phonological issue rather than motoric. However, McAllister Byun (2012) suggests that the ULG patterns reported by Gibbon (1999) are explained by constraints in independent movements of the tongue and jaw. Early in life, children are unable to decouple tongue and jaw movement, instead moving them as one unit (Davis & MacNeilage, 1995). This “move as unit” constraint particularly applies to more ballistic movements, making gestural magnitude important. Since prosodically strong contexts, such as syllable onsets, have been shown to have a greater gestural magnitude, this makes them particularly vulnerable to the “move as unit” constraint. This results in attempts at velar stops, which are more likely to involve whole movement of the tongue toward, and over-contact with, the palate. In turn, this “move as unit” constraint becomes encoded in the child’s phonology, making velar fronting an error that exemplifies the importance of considering both phonetic and phonological explanations of SSDs together.

So, children with persistent velar fronting may lack sufficiently mature speech-motor control to enable them to move the coronal and dorsal parts of the tongue independently in all contexts, or to adequately decouple tongue and jaw movement. While precise models remain to be formulated, the resulting gestural error from these causes can be similar: the lack of a functional coronal/dorsal distinction.

Another topic that has not been explored much in the literature, and that is also essential for the building of predictive models, is the detail of how these children overcome their difficulties and ultimately develop successful alveolar–velar contrasts. While naturally resolving speech-motor delay might be an explanation in some children who develop the contrast slightly later and without intervention, we argue that insights can be gained from the small group of children

with more persistent velar fronting who have not responded to phonological intervention. Few studies have attempted to look at the processes involved in acquiring a new articulatory gesture, in this case a dorsal gesture that creates the percept of a velar stop. For obvious reasons, most studies report the functional percentage of target consonants correct at various time points in intervention. However, given how ULGs can lead to difficulties with transcription, this does not necessarily give enough insight into the processes by which children learn a new gesture. Moreover, in traditional phonological interventions for velar fronting, it is normally presumed that children are stimutable for velars at the outset of intervention, and hence the focus is on improving phonological knowledge (Dodd et al., 2018). We might, therefore, predict that if the underlying cause of velar fronting is phonological, we would see velar gestures at the beginning of intervention that are similar to those produced by typical children. That is, the problem is not with producing an appropriate gesture but with applying the phonological rules required to use the gesture in multiple contexts. Conversely, if children (also) have a motoric difficulty producing velar gestures, we might expect to see a gradient or variable acquisition of the new gesture consistent with the view that learning a new motor gesture is slow and inconsistent (Weaver, 2015).

Before investigating how children with disordered productions learn new, more adultlike gestures, it is important to determine what TD children of different ages do, to differentiate their alveolar and velar consonants at the articulatory level. Few studies have reported data on TD children. Gibbon (1999) summarized EPG data from 24 typical children across six studies to demonstrate that typical tongue–palate contact for /t/ and /k/ involve small amounts of contact in the lateral margins and alveolar region for /t/ and contact in (or beyond) the velar region for /k/. However, these norms are from multiple studies using multiple different speech materials. In a larger study by Cheng et al. (2007), data on /t/ and /k/ productions were collected across four different ages of speakers: 6–7, 8–11, 12–17, and over 18 years. EPG patterns show, that in the youngest age group of children, /t/ was produced with greater amounts of tongue–palate contact compared to the older three age groups, whereas contact was relatively stable for /k/ productions. However, for most speakers, the EPG plate did not capture the velar closure for /k/ as plates typically end at the juncture of the hard and soft palates (Cleland & Preston, 2021). It is also important to note that, although more electrodes were activated in the youngest age group, this does not necessarily imply a greater surface area of tongue–palate contact in absolute terms as EPG palate displays are normalized—that is contacts are closer together in speakers with smaller hard palates. While these norms are useful for clinicians wanting to use EPG, they do not readily translate to other articulatory techniques. Moreover, the use of EPG is declining in the clinic, while the use of ultrasound visual biofeedback is increasing (Sugden et al., 2019). It would, therefore, be useful to have norms using UTI since they can be applied to larger groups of children and to the

increasing numbers of clinical studies using this technique for biofeedback.

### *UTI Studies of the Alveolar–Velar Contrast*

UTI is becoming an increasingly popular technique in the phonetics laboratory for measuring tongue shape and movement and in the speech therapy clinic for visual biofeedback (Sugden et al., 2019). There are essentially two different ways of analyzing ultrasound data, dependent on the recording setup. The first involves looking at individual tongue splines and taking ratio measures within a single speech sample, using the ratios as the basis for further analysis. This method is more resistant to changes in probe location due to translation and rotation within the midsagittal plane and is, therefore, useful if data are collected without stabilizing the probe relative to the head (Zharkova, 2013).

An important ratio measure for differentiating alveolars and velars is the Dorsum Excursion Index (DEI; Zharkova, 2013). DEI is a measure of the magnitude of excursion of the tongue dorsum, relative to the front–back extension of the tongue. Higher DEI values indicate more dorsum excursion, and therefore, we expect a typical /k/ articulation to have a higher DEI than a typical /t/. There are no studies using this measure on sizable numbers of TD children, but it is possible to extrapolate from Zharkova (2019), which presents data from six typical adults. This study shows that DEI reliably distinguishes /k/ and /t/ in an open vowel context, for example, /a\_a/. However, this measure (and others presented in the article) could not differentiate between /k/ and /t/ in high-vowel contexts, despite these consonants having different primary constrictions. The largest study of children of which we are aware analyzes the consonant production of four Farsi-speaking children. Baghban et al. (2020) found that DEI and another measure, the Tongue Constraint Position Index (Zharkova, 2019), both successfully distinguished the children's /k/ from /t/. However, this again does not easily provide us with norms with which to compare children with SSDs.

To our knowledge, only one study has applied DEI to children with SSDs. McAllister Byun et al. (2016) applied the measure to two children with velar fronting and two children who had already developed the contrast. This study was specifically looking for evidence of covert contrasts in the speech of the children who still displayed the neutralization. They were, therefore, expecting to see small differences in DEI between /k/ and /t/ attempts in children with a perceptual neutralization. Results showed that DEI was able to reliably separate /t/ and /k/ articulations in both front (“key vs. tea”) and back vowel contexts in the children with the overt contrast and could detect a covert contrast in one of the two children with velar fronting.

A second method for analyzing ultrasound data involves direct comparison of tongue curves. In this method, the probe must be stabilized relative to the head using a headset (Scobbie et al., 2018) or there must be a post hoc correction for probe movement, for example, using optical

tracking (Whalen et al., 2004). Stabilization and/or correction allows the researcher to measure differences between tongue shapes directly from a larger data set, and a headset prevents even within-syllable probe dislocation from the midsagittal plane, which makes the ratios inaccurate. Melo et al. (2017) used this method to collect data on alveolar–velar productions in 15 TD Brazilian–Portuguese children aged 4;7–7;5 and 20 adult speakers. They first averaged multiple repetitions of productions of /k/ and /t/ in open vowel contexts and then compared them using the inbuilt *t*-test function in Articulate Assistant Advanced software (Articulate Instruments Ltd, 2014). This *t*-test function uses a fan-shaped measurement space that aligns with the fan-shaped imaging area, which emanates from an ultrasound probe. Using the probe as the origin, a set number of fan-lines is drawn radially from the probe. (The number varies depending on the ultrasound system used. Here, it is 42; see Figure 1.) Melo et al. (2017) report the number of significant differences in radial fanlines in two separate regions roughly comprising the front and back of the tongue. The *t* test reliably distinguished /t/ and /k/ in TD children. Qualitative visual inspection of the tongue curves suggested that compared to adults, children showed less differentiation between tip and dorsum gestures and greater variability in multiple repetitions.

Two studies have used *t* tests of articulatory data to make observations about alveolar–velar contrasts in English-speaking children with persistent velar fronting. In the first study, Cleland et al. (2015) compared attempts at /k/ before (all transcribed as [t]) and after (all transcribed as [k]) intervention with ultrasound visual biofeedback in four children with persistent velar fronting. These tongue curves showed a significant *t*-test result along at least six out of 42 adjacent fanlines. Moreover, Cleland et al. (2015) also reported the average and maximum spatial radial dorsal difference between [t] and [k], with maximum differences ranging from

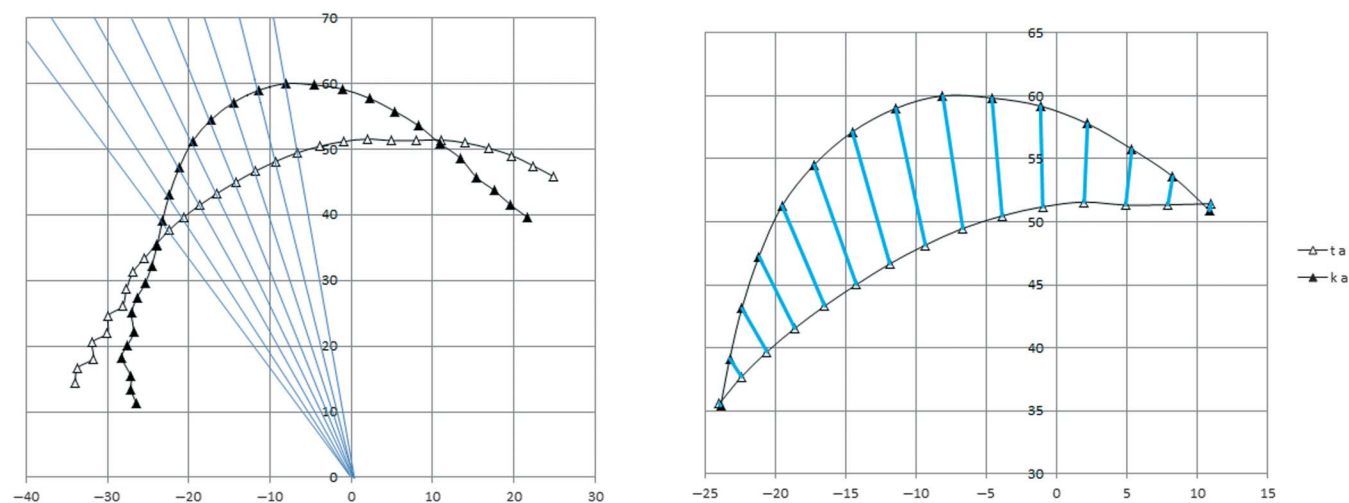
6.5 to 11 mm. In a follow-up study involving the same children, plus an additional three children with persistent velar fronting, Cleland et al. (2017) used the *t*-test measure to demonstrate a lack of any covert contrast between /t/ and /k/ pre-intervention, signified by nonsignificant *t*-test results supported qualitatively by near-identical tongue shapes for both consonants. However, visual inspection of tongue-shaped data for one child in the study revealed that she produced /t/ and /k/ with a variety of tongue shapes including ULGs. Neither study compared the children with SSD to TD children.

## Purpose

Two studies were carried out. First, a study of TD children aimed to provide typical development norms for the alveolar–velar distinction. Second, a study of children with a history of persistent velar fronting aimed to illustrate the utility of these norms. In Study 1, we used a measure that distinguishes /k/ and /t/ by quantifying the magnitude of this distinction in absolute spatial terms (mm of linear dorsal difference). As the measure was absolute, we hypothesized that it would correlate positively with chronological age because older children are expected to have larger vocal tracts. We also hypothesized that the norms would differ by vowel context due to coarticulatory effects. We, therefore, expected a larger difference between alveolars and velars in low or back vowel contexts than in a high, front vowel context where velars are normally fronted toward the palatal region (Frisch & Wodzinski, 2016).

Study 2 then applied these norms to illustrate the utility of this measure for classifying and quantifying the speech of children with a history of persistent velar fronting as they develop the contrast longitudinally. In this study, we also applied the *t*-test measure described in Melo et al. (2017) and Cleland et al. (2015) to determine whether children

**Figure 1.** Left: Example /t/ (white triangles) and /k/ (black triangles) tongue splines with spline knots. Anterior is to the right. Blue lines show the radial fanlines with the probe as origin. Right: KT crescent showing distances between knots along radial fanlines.





with persistent velar fronting achieve statistically significant differences in tongue shape for /t/ and /k/ before, during, and after intervention. We hypothesized that prior to intervention, children with a history of velar fronting would show values near zero in our new measure, coupled with nonsignificant *t*-test results comparing tongue-shape splines for /t/ and /k/ that is, a complete merger between /t/ and /k/ (Cleland et al., 2017). Postintervention, we expected that children would show values similar to the norms for typical children, and significant *t*-test results, suggesting successful acquisition of the contrast. During intervention, we hypothesized that a sudden change from values near zero to values in the normal range would suggest a categorical shift in production of velars, perhaps indicating a problem of phonological origin, whereas a gradual increase in values, or variable values, would likely indicate motor difficulties. In both of these contrasting possibilities, we expected to find significant *t*-test results as children began to establish a perceptible contrast between /t/ and /k/, even if this contrast was realized in an articulatorily abnormal manner. Both studies involved analysis of data freely available in the UltraSuite corpus (Eshky et al., 2018), an Open Data approach that enables future comparison with other measures or replication by other researchers.

## Study 1, Norms for TD Children

### Method

Participants were 30 TD English-speaking children reported in the UltraSuite Corpus (Eshky et al., 2018). Children were recruited via advertisements within Queen Margaret University, Edinburgh. Ethical approval for this study was provided by the same university's institutional review board. The children were aged 5;8–12;10 ( $M = 9.52$ ;  $SD = 2.04$ ), and as such, all had acquired velars. The children were screened for speech disorders using the Diagnostic Evaluation of Articulation and Phonology (Dodd et al., 2002) and language delay using the British Picture Vocabulary Scale–Second Edition (BPVS; Dunn et al., 1997). One child showed evidence of speech delay on the Diagnostic Evaluation of Articulation and Phonology and language delay on the BPVS, and another child showed evidence of language delay on the BPVS. Both children were retained in the analysis as they had acquired velar and alveolar stops.

### Materials

The UltraSuite corpus of TD children contains a word list designed to sample all consonants and vowels of Scottish English. Consonants were elicited by imitation of an audio prompt comprising each consonant in isolation followed by three intervocalic contexts, /aCa iCi oCo/, with initial stress. These three vowels represent the corner vowels of the Scottish English vowel space. We chose /o/ (a monophthong typically higher than cardinal [o]) as the highest–backest vowel in Scottish English since /u/ is fronted and lowered

(Scobbie et al., 2012). The low vowel /a/ is slightly more central than cardinal [a] and is used for the lexical sets TRAP, BATH, and PALM (Wells, 1982). We, therefore, analyzed [aka iki oko] and [ata iti oto], none of which we regard as real words of English, but all of which are phototactically permissible. One token of each consonant in each vowel context was produced and is analyzed here, giving 180 tokens in total for the analysis, 30 in each condition.

### Ultrasound Recording

The UltraSuite corpus data were collected using a high-speed cineloop system with audio synchronization and a probe-stabilizing headset (Scobbie; Scobbie et al., 2018). This allowed us to compare tongue shape for /t/ and /k/ directly. The headset was fitted in such a way that the mandible and hyoid shadows were symmetrical on the image, thus ensuring that as much of the tongue was visible as possible.

Ultrasound data were acquired using an Ultrasonix SonixRP machine remotely controlled via Ethernet from a PC running Articulate Assistant Advanced software (Versions 2.14 to 2.16; Articulate Instruments Ltd, 2014), which internally synchronized the ultrasound and audio data. The echo return data were recorded at 121 frames per second, that is, 8 ms per frame, with a 135° field of view in the mid-sagittal plane. Simultaneous acoustic recordings were also made, using an Audio-Technica 803D clip-on microphone sampling at 22050 Hz.

### Analysis

Using AAA Version 2.16 software (Articulate Instruments, 2012), /t/ and /k/ segments were annotated at the point of maximum constriction. If any ULGs had been present in the data, the point of maximum constriction would have captured this gesture; however, these types of gestures were not found in the TD children. The corresponding ultrasound frame was then selected, and a spline indicating the tongue surface fitted to the image using the semi-automatic edge-detection function in AAA software. For each speaker, the difference between /t/ and /k/ was computed as the speaker's own overlap of /k/ and /t/ in the midsagittal plane, in each vowel context independently.

In articulation, tongue-surface shapes of correctly produced /k/ and /t/ overlap spatially. The extra dorsal constriction for /k/ (together with a lowered blade) produces a crescent-shaped dorsal overlap (see Figure 1). The crescent is bounded front and back by two crossover points that define the tips of the crescent. These crossovers indicate that, in two particular points in the plane, /k/ and /t/ happen to have the same location. However, a different part of the tongue surface is responsible in each case, and this is a side effect of the different primary places of constriction.

In order to determine the degree of dorsal separation between /t/ and /k/, we calculated both the area of the crescent and its maximum linear depth, which is the maximum radial difference of /k/ minus /t/, based on the ultrasound probe as origin of the measurement space. Polar coordinates were used rather than Cartesian because, in AAA, a fan-shaped grid with 42 radii is used both to semi-automatically

fit splines to the image and to measure, average, or export them via (a subset of the available) 42 spline knots. AAA, therefore, easily provides the distance of each spline from the fan's origin (i.e., the nominal center of the probe) along each radius (which are numbered).

### KTmax Measure

For each radius, we exported two spline knot locations to MS Excel as polar coordinates (i.e., distance from the origin at a given fan angle). For each vowel environment, the radial differences between /k/ and /t/ are simply the differences between the two knots on each fanline in which /k/ is further from the probe than /t/, which together generate a crescent shape, as noted above. We can then use the single maximum radial difference in mm between /k/ and /t/ within this crescent shape as a measure of the degree of differentiation between the stops in the dorsal region.

### KT Crescent Area

In addition to the maximum radial difference, we calculated the midsagittal area of the crescent in each vowel environment. Given the 135° field of view of the original image, onto which the fan was fitted, the angle between each equally spaced radius was (135/42) 3.21°. Figure 1 (left) shows a typical /t/ and /k/ with triangles representing the spline knots.

The midsagittal area of the difference between the splines was operationalized as the sum of a number of annular sectors (see Figure 1, right). These regular shapes can be conceptualized either as the difference between equal-angle sectors of two concentric circles, or as a sector of an annulus (a ring). In this example, there are 13 radii within the irregular crescent (see Figure 1, right), so 13 annular sectors can be summed to obtain its area. Sectors centered on the radii about each other, but do overlap the inner and outer splines. Figure 2 shows how this error nevertheless lets us approximate the area between /k/ and /t/. While one corner overestimates the area by overlapping the spline, the corresponding corner underestimates, by a similar amount. In this case, each annual sector is 0.89% of a hypothetical complete annulus (3.21°/360). Each fanline radius within the KT crescent therefore bifurcates an annular sector,

which extends symmetrically on either side, clockwise and anticlockwise. The area (A) is based on  $\Theta$  (the between-radii angle in degrees, here 3.214°), K (radial distance to the further knot, belonging to /k/), and T (radial distance to the closer knot, belonging to /t/):

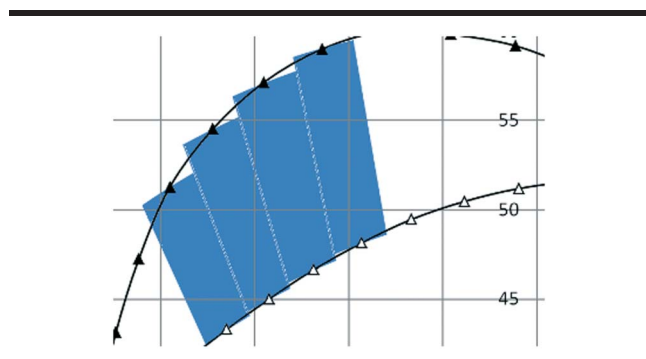
$$A = (\theta/360) \times \pi(K^2 - T^2). \quad (1)$$

For some speakers, the KT crescent did not have two perfectly identifiable crossover points; however, we included these data because the area of missing annular sectors involved was very small at the tips of the crescent. Figure 3 shows example data with two identifiable crossing points (right), missing anterior crossing points (left), and missing posterior crossing points (middle). The following assumptions were used: First, the perfect case was when there were crossovers between the /k/ spline and /t/ spline to define the two crescent tips. Since area was calculated on contiguous radii within the crescent for which  $K > T$ , there was, by definition, always some overshoot or undershoot of the actual crossing point unless it was *perfectly* coincident on two knots. This did not ever occur. If the splines came close, but no crossover occurred due to missing data at the end of the tongue curve, then the area was computed missing the tip of the crescent, thus slightly underestimating the area. This was unusual and occurred on only 23/180 occasions, mostly at the anterior crossing point for /a/ (see Figure 3, left) and /i/ and at the posterior crossing point for /o/ (see Figure 3, middle). If the splines came close yet stayed nearly parallel without crossing, a judgment was made to terminate the crescent *as if* there had been a crossing point. It was defined to be where the splines first reached the stable close approximation (e.g., around 1–2 mm). This was uncommon ( $n = 12$ ) and affected only the tongue root. The number of times a root spline crossed then recrossed was also rare ( $n = 6$ ).

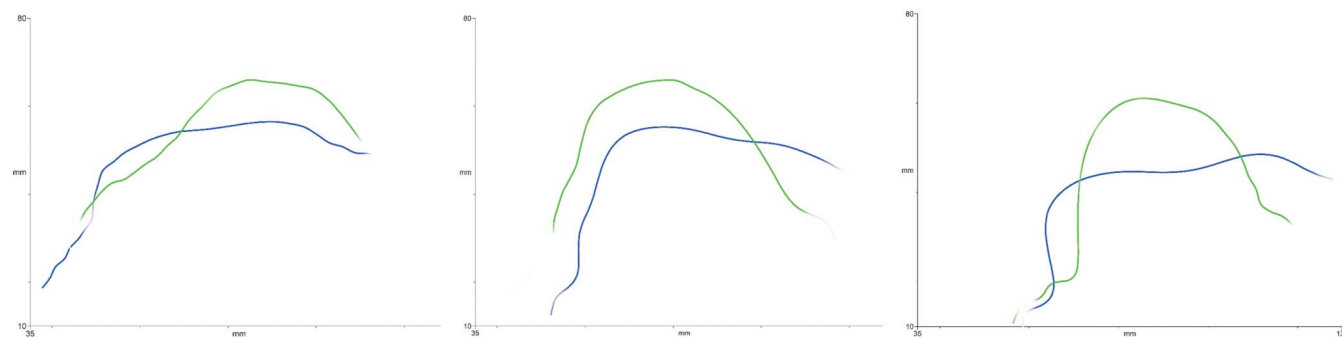
A number of subsidiary analyses were performed to evaluate these measures. We tested our expectation that KTmax and KT Crescent Area would be positively correlated since the crescent shapes were relatively regular. We also analyzed whether these measures correlated positively with the age of speakers as we expected older children to have larger vocal tracts. Interrater reliability was calculated by a second annotator (the first author) relabeling a randomly selected 20% of the data and recalculating KTMax. Interrater reliability was excellent, the average measure ICC was .990 with a 95% confidence interval from .956 to .996,  $F(23,23) = 138.31$ ,  $p < .001$ . In absolute terms, KTMax between raters was on average within 1 mm ( $M = \pm 0.75$  mm,  $SD = 0.64$ ).

In order to create a visual representation of the norms that might be used in intervention as target tongue shapes, we averaged rotationally normalized splines. Normalization was based on the maximum constriction for /aka/ in order to compare both the size and location of the /k/-/t/ difference in space and across different vowel contexts. In other words, three allophones of /k/ and the three of /t/ were averaged across speakers by aligning each speaker along the

**Figure 2.** Annular sectors within the KT crescent.



**Figure 3.** Example /t/ (blue) and /k/ (green) tongue splines from typically developing children. Left to right: missing anterior crossing points; missing posterior crossing points; and visible anterior and posterior crossing points. Anterior to the right.



maximum radial distance to the tongue surface in the /k/ of /aka/.

## Study 1 Results

### KTmax Measure

Table 1 shows the maximum radial difference of /k/ minus /t/ (KTmax) for each of the vowel contexts. Both /a/ and /o/ contexts have similar KTmax of around 12 mm, whereas the difference is only around 7 mm in an /i/ context. The effect of vowel was significant,  $F(2,28) = 25.34$ ,  $p < .0001$ , as expected.

### KT Crescent Area

The crescent areas for each of the vowel contexts are also shown in Table 1. The mean area of the crescent-shaped velar difference between /k/ and /t/ varied by vowel context,  $F(2,87) = 17.34$ ,  $p < .0001$ , with the /i/ context having about half of the area of the /a/ and /o/ contexts. The “width” of the KT crescents, that is, the number of radii between crossing points, was fairly consistent. It spanned roughly 45° relative to the probe origin, a significant proportion of the vocal tract as imaged from that location. The difference in area in the /i\_i/ context is due to the palatalization of both consonants resulting in a smaller difference in the maximum radial difference between palatalized /k/ and /t/.

**Table 1.** Quantitative aspects of each typically developing children’s KT crescents.

Vowel	Area (mm <sup>2</sup> )			# radii (each 3.2°)			Max rad diff (mm)		
	a	i	o	a	i	o	a	i	o
M	294	164	272	13.9	12.8	13.9	11.9	7.5	12.1
SD	83	82	90	3.1	3.1	3.2	3.0	3.4	3.8
Min	159	43	52	9	6	5	7.3	3.3	5.9
Max	450	405	554	19	19	21	18.0	16.0	22.0
Mode				13	14	13			

*Note.* Max rad diff = the maximum radial difference between /k/ and /t/ in mm.

## Correlations

There was a strong correlation between KTmax and KT crescent area ( $r = .791$ ,  $p < .0001$ ) at a linear rate of 1:23. Contrary to our expectations, there were no significant correlations between KTmax or KT crescent and chronological age in any vowel environment.

## Visualized Norms

Figure 4 shows /t/ and /k/ splines across all three vowel contexts averaged across all 30 participants. These visualized norms illustrate the influence each context has on the difference between /k/ and /t/ and are intended as useful comparisons for disordered speech or as useful target tongue shapes for intervention.

## Study 2: Acquisition of Velars in Children With SSD

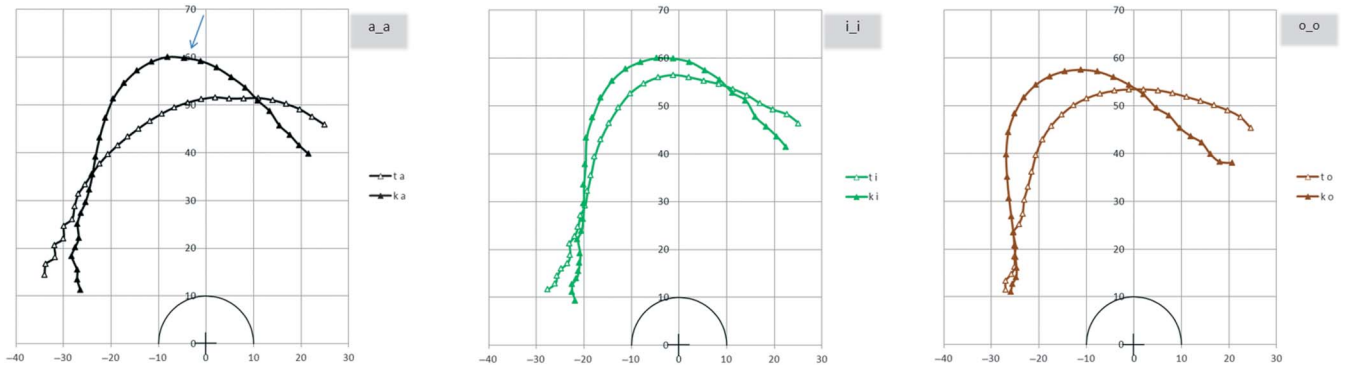
### Method

Participants in this study were five children with persistent SSD: 01F, 03F, 04M, 15M, and 17M. Data are available in the UltraSuite Corpus (Eshky et al., 2018). Ethical approval was provided by the National Health Service Ethics Committee (the local health board). All of the children received between nine and 12 sessions of ultrasound visual biofeedback for the treatment of velar fronting. The treatment and its outcomes are reported in detail in Cleland et al. (2019). We selected here a subset of children who were treated for velar fronting, irrespective of SSD subtype, and who showed, on the basis of phonetic transcription, some improvement in acquisition of velars during intervention. Participants 01F, 03F, and 04M were also reported in Cleland et al. (2017), in which they were shown to produce no covert contrast between /k/ and /t/ pre-intervention, evidenced by identical tongue shapes for /k/ and /t/ attempts. Table 2 presents the children’s demographic details and a summary of their intervention progress.

## Materials

In this clinical study, we were primarily interested in the children’s acquisition of correct-sounding velars over

**Figure 4.** Averaged /t/ (white triangles) and /k/ (black triangles) tongue splines in different vowel contexts. Left to right: /aCa/, /iCi/, and /o\_o/. Anterior is to the right.



time. We expected the children to begin intervention with KTM<sub>ax</sub> scores near zero and to end intervention with KTM<sub>ax</sub> scores similar to the typical children. During intervention, we hypothesized that a categorical shift from KTM<sub>ax</sub>  $\approx$  zero to KTM<sub>ax</sub> within the normal range for typical children would suggest a phonological basis to the error, whereas a more gradual or variable change in KTM<sub>ax</sub> would align with the view of Gibbon (1999) and McAllister Byun (2012) that velar fronting is an error due to motor constraints. We therefore focus here primarily on the in-session recordings made during the intervention sessions. These were recorded with an identical setup to Study 1, including headset stabilization. However, there were no standard speech materials in each session. Instead, the treating clinician started each session by recording the participant attempting 10 productions of velars in single syllables, words, or sentences, depending on their progress to date in intervention (see Cleland et al., 2019, for further details).

Further recordings were made during the intervention sessions when the clinician judged that the child was making progress in an ad hoc manner. Some of these recordings were during the prepractice phase of intervention (Maas et al., 2008) where children were being given specific instructions on how to produce an articulatory gesture and specific feedback about the closeness and acceptability of their attempts to prepare the children for the practice phase of intervention. Some children achieved near articulations

(e.g., uvular or velar fricatives) during prepractice, but did not move to the practice phase of intervention during that particular session. We included these dorsal articulations in our analyses if there were no correct velar stops. Given that multiple repetitions were available, and Melo et al. (2017) had shown the AAA *t* test to be a useful measure of differentiating /t/ and /k/, we used multiple repetitions of alveolars and dorsal articulations in Study 2. We hand-searched each session for data that could be used to define a dorsal crescent against an alveolar baseline using the following criteria, bearing in mind that, unlike TD children, there would not always be simple [k] for /k/ and [t] for /t/ in the sample: (a) There were at least three (and up to 10) impressionistically correct or near attempts at the dorsal consonant in an open vowel context. (b) The tokens selected were judged (on impressionistic transcription grounds) to be the best of the child’s attempt at a dorsal articulation, even if it was a uvular articulation or a ULG. (c) A comparison alveolar in the same vowel context was available. We used any available alveolar in the same vowel context, but preferred /t/. We assumed, in particular, that productions of /n/ would be similar enough in the midsagittal view to /t/ (Gibbon et al., 2007). For some sessions, especially sessions early in intervention, an alveolar target was not available at all. We therefore chose to use unambiguously alveolar phonetic productions of attempts at /k/ as comparisons, if these were realized as [t]. Incorrect /k/ was, therefore, a proxy

**Table 2.** Demographic details of children with speech sound disorder.

Participant	Sex	Age	SSD subtype*	% Velars correct		
				Preintervention	Postintervention	6 weeks postintervention
01F	F	8;8	Inconsistent speech disorder	5	22	/
03F	F	10;11	Childhood apraxia of speech	2	29	60
04M	M	7;2	Phonological delay	2	55	61
15M	M	6;1	Phonological delay	10	42	82
17M	M	13;2	Phonological delay	2	18	17

\*SSD subtypes were defined by the referring clinician.



for our definition of alveolar, to compare with nonalveolar productions of /k/ and to provide a basis of comparison with other sessions when nonalveolar productions of /k/ could be compared to true alveolars.

Table 3 details the speech materials used for each child in each session alongside an International Phonetic Alphabet transcription of the child's attempts, and the number of repetitions used. Where the target was produced within a real word, the word is given to the right of the context or comparison context, otherwise tokens were produced in

consonant–vowel/vowel–consonant (CV/VC) nonwords or in isolation. Some sessions were missing data due to a recording failure or a difficulty finding comparison data.

For comparison with pretherapy productions of velars, we also included the productions of /k/ and /t/ in untreated word lists before, immediately post, and 6 weeks postintervention. These word lists comprised productions of /t/ and /k/ in a wide variety of contexts in untreated word lists. They are therefore not directly comparable to the “best attempts” in intervention, which were scaffolded

**Table 3.** Speech materials analyzed from each session.

Participant	Session	Context <sup>a</sup>	IPA	Reps	Comparison <sup>a</sup>	
01F	1	ok	t	3	k attempt	ko
	2	ok	q	10	k attempt	ko
	3	xk	q	3	k attempt	ko
	4	xk	x	2	k attempt	ko
	5	xk	q	4	k attempt	ko
	6	ok	q	8	k attempt	ko
	7	xk	fɪjk	7	k attempt	ko
	9	ka	k	10	t	tap
	10	ka	k	10	n	cone
03F	1	ok	k	10	k attempt	ko
	2	ko	k	10	k attempt	ko
	3	ko	k	10	n	cone
	4	ka	k	10	t	kite
	5	ka	k	10	t	tacky
	6	ka	k	10	t	tacky
	7	ak	k	10	t	tacky
	8	ko cone	k	5	n	cone
	9	ko cone	k	10	n	cone
	10	ko cone	k	10	n	cone
04M	1	ko	k	9	k	ko
	2	ko cone	k	9	n	cone
	3	ka cap	q	8	t	tacky
	4	ak tacky	k	10	t	tacky
	5	ko cne	q	12	n	cone
	6	ka calculator	q	10	t	tap
	7	kai kite	q	10	n	playin'
	8	ak back	q	5	t	eats
	10	ka car	q	10	t	tap
15M	1	ok	χ	7	k attempt	ko
	2	ok	χ	10	k attempt	ko
	3	ok	χ	10	k attempt	ko
	4	oɪj	N	3	k attempt	ko
	5	oɪj	N	10	k attempt	ko
	7	ok	q	13	k attempt	ko
	9	ok	q	10	n	ko
	10	ka cap	c	10	t	kite
	11	ko cone	k	6	n	cone
	12	ko cone	k	4	n	cone
17M	1	x	h	9	k attempt	ko
	2	xk	χq	10	k attempt	ko
	3	ko	q	10	k attempt	ko
	4	oko	q	9	n	cone
	5	ko cone	k	10	n	cone
	6	ka car	k	10	t	kite
	7	ak tacky	k	10	t	tacky
	8	ak tacky	k	10	t	tacky
	9	ka calculator	k	6	t	calculator

<sup>a</sup>Where the articulation was produced within a real word this is given to the right of the “context” or “comparison.” Blank cells indicate the dorsal articulation was produced in isolation, consonant–vowel (CV), vowel–consonant (VC), or consonant–consonant (CC) nonword context. IPA = International Phonetic Alphabet; F = female; M = male.

by the treating clinician. Instead, they are a measure of generalization. We selected all productions of /t/ and all productions of /k/ transcribed as acceptable productions in a low vowel context. Any incorrect productions of /k/ were transcribed as [t] and clearly alveolar.

## Analysis

We annotated the alveolar and dorsal articulations at the point of maximum constriction using AAA in a similar manner to Study 1. Again, using the maximum point of constriction, rather than the release, means that any ULGs were captured. Note, not all attempts were oral plosives—early in intervention, some children produced velar or uvular fricatives. An additional step to Study 1 was added in that, within each session, we first averaged the multiple attempts at the velars and, separately, comparison alveolars. By analyzing multiple repetitions, we were also able to compute *t* tests using the inbuilt function in AAA software to determine whether each child was producing a reliable statistical difference between alveolar and velar attempts. In this type of *t* test, significance is tested radially along each of the 42 fanlines. Our threshold for reporting significant difference between means is a minimum of six adjacent radii where the mean difference is significant at  $p < .05$  (Cleland et al., 2015). These six adjacent radii are over a contiguous region of the tongue surface (approximately 2–3 cm of surface), reflecting the fact that the adjacent parts of the tongue and their distance from the origin of the fan-grid (i.e., the center of the probe) are not independent.

We also computed KTmax to determine the magnitude of that difference (significant or not) and as a method of validating the new measure. Since we did not have consistent vowel contexts in each session, we also computed a KTmaxnorm measure by dividing KTmax by the value of the TD mean for the relevant vowel used (either /a/ or /o/). This allowed us to report ratios where numbers close to zero represent no difference between alveolars and velars and close to one represent values close to the norm.

Since TD children vary, we report whether children with SSD were within  $\pm 1$  *SD* of the mean and within the full range of all TD children. We did not account for age since there was no correlation between age and KTmax in the TD children. Since the children produced a variety of unusual tongue shapes (see illustrations below), we were less often able to reliably compute the “malformed” crescent area than anticipated, as crossing points were often not available. For simplicity, we therefore report, as follows, using the more tolerant (highly correlated) linear measure: (a) the *t*-test results for each child, at each time point; (b) KTmaxnorm at each time point for each child; (c) narrow phonetic transcriptions of the velar attempts; and (d) qualitative illustrations of some unusual tongue shapes.

## Study 2 Results

A total of 48 intervention sessions across the five participants had useable data. All children had pre- and postintervention data available, and all children except 01F had

6-week postintervention data available. (She did not return for the final recording.) Across all intervention sessions, there was an average of 8.57 repetitions of velar consonants available for analysis ( $SD = 2.54$ , range: 3–10). Figures 5a–5e show the individual data for each child. KTmax norm is plotted across time. We also report the *t*-test results (average *p* values across at least six adjacent fanlines) at each time point for convenience on the same figures. Phonetic transcriptions are given next to each data point. It should be noted that, although there were multiple repetitions of the dorsal articulation, the transcriptions are the same for each. This occurred because the items were produced during blocked practice; other productions in the same session often varied. Each figure shows in gray bands  $\pm 1$  *SD* of the norms and the full range of values from TD children. Results will be discussed for each child individually and then summarized.

### 01F Results

Pre-intervention 01F had KTmaxnorm values near zero and a nonsignificant *t* test as predicted, indicating no difference between /t/ and /k/ productions, which were [t]-like. Postintervention, 01F showed values within 1 *SD* of the normal range, and /k/ was transcribed as [k]. It is worth noting that, at this time point, 01F continued to make many categorical errors and was variable across the phoneme boundary: only productions involving dorsal raising are, therefore, included in the analysis here.

During intervention, 01F showed a range of KTmaxnorm values, including very high values, taken from tokens transcribed as uvular productions. This range of values is consistent with the notion that, for this child, velar fronting is likely due to a motoric problem. Time Point 7 is interesting in that a low KTmaxnorm was obtained, yet the transcription of that seems at odds with the value. Figure 6 shows why: 01F produced /k/ with a typical tongue shape, but alveolar stops showed wide variability, including ULGs. This patterning has been explored further in a case study—see Cleland et al. (2017). This illustrates that the KTnorm measure is a measure of separation between /k/ and /t/, not a measure of correctness per se. This is especially important to remember when addressing productions by children with motor speech disorders who may present with abnormal tongue shapes.

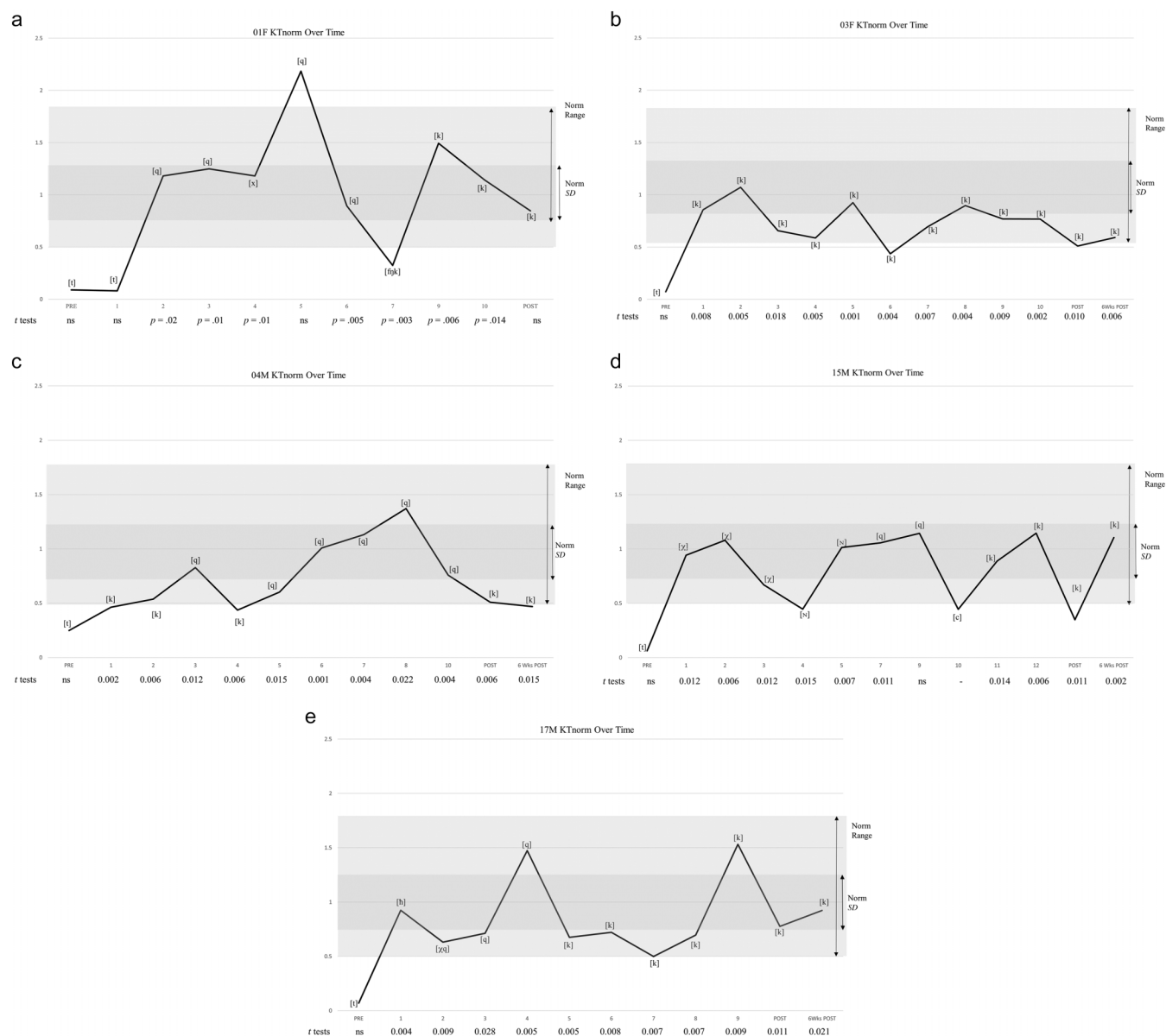
### 03F Results

Pre-intervention 03F also showed values near zero and a nonsignificant *t* test. Postintervention and during intervention, she was able to achieve productions of /k/, which were transcribed as [k] and which were significantly different to /t/, as shown by the *t*-test results. KTmaxnorm values were mostly within the range of those produced by TD children, although for the most part they were not within 1 *SD* of the mean.

### 04M Results

04M also produced values close to zero and nonsignificant *t*-test results pre-intervention and values close to

**Figure 5.** (a–e): KTmaxnorm over time for children with persistent velar fronting. Normalized KTmax values are shown on the y-axis; time points are on the x-axis. Phonetic transcriptions for each time point are given next to data points. Dark gray shading shows the mean value for typically developing children  $\pm 1$  SD; light gray shading shows the full range for typical children. Average  $t$ -test results for each time point are given under the x-axis.



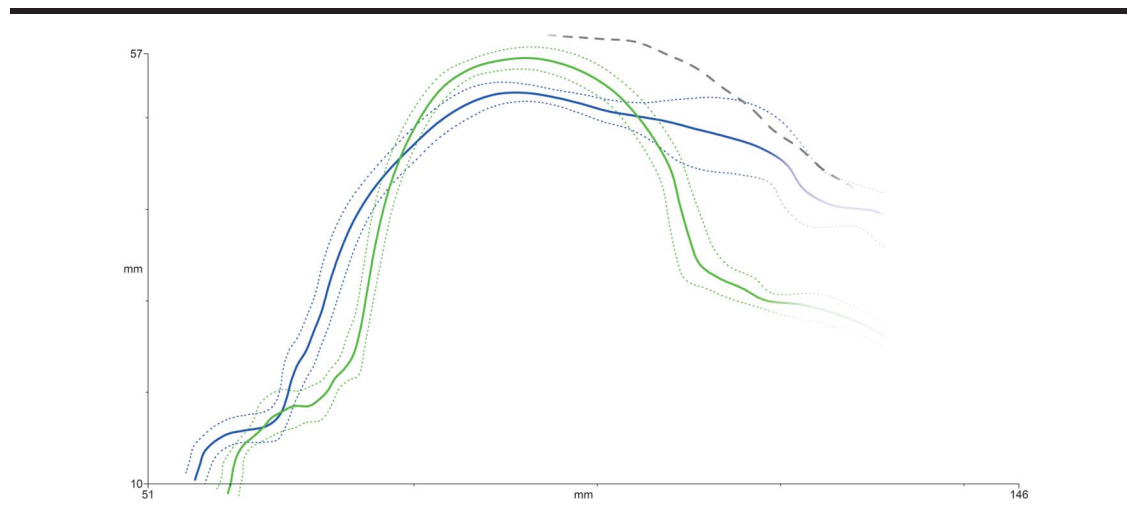
the normal range postintervention. During intervention, many of his productions were transcribed as uvular, but were within the normal quantitative range. Productions that were transcribed as a correct [k] had quite low KT values, suggesting that even low values may indicate an appropriate target, at least for this child.

### 15M Results

15M again produced a radial difference near zero pre-intervention and within the normal range postintervention. Again,  $t$ -test results followed the expected pattern. During

intervention, uvular productions were common. In one session (10), he produced palatal stops. Both types of distortion suggest that he had not yet consistently mastered the gesture. (For Session 10, only one comparison alveolar was available, hence no  $t$ -test result). Session 4 is of interest: A low KTmaxnorm value was obtained from a uvular production. Figure 7 shows that the dorsal articulation was clearly retracted to uvular and that there was a large anterior difference between attempts at /k/ and /t/. A dorsal measure does not, of course, directly incorporate an anterior difference; however, the small dorsal value does suggest

**Figure 6.** Example average tongue splines for /k/ (green) and /t/ (blue) showing a raised production of /t/. Anterior is to the right.



an abnormal /k/ or /t/ gesture (or both). A visual inspection of the tongue shapes is required to determine the exact nature of the unusual gesture(s).

### 17M Results

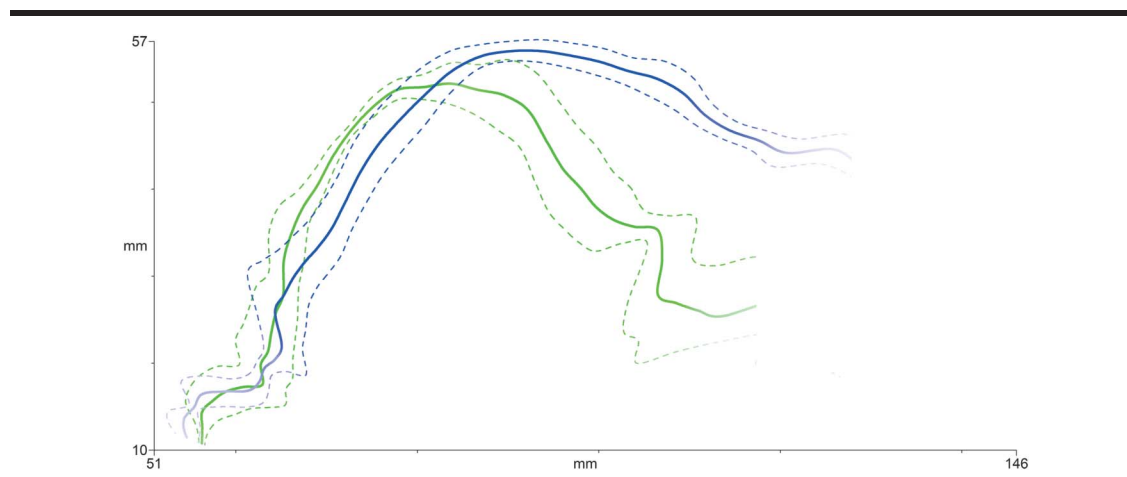
Again, 17M produced values near zero and nonsignificant *t*-test results pre-intervention and moved within the normal range postintervention. During intervention, he produced a variety of different values, beginning intervention with pharyngeal or uvular articulations that nonetheless were characterized by KTmaxnorm values in the range for TD children.

### Study 2 Results Summary

All five children began intervention with KTmaxnorm values near zero and nonsignificant *t*-test results comparing tongue curves for /k/ and /t/, confirming our hypothesis. The

mean KTmax across the group pre-intervention was only 1.29 mm ( $SD = 0.94$ ). This suggests that, for all children, there was a complete merger of /k/ and /t/ prior to intervention, as shown in Cleland et al. (2017). During intervention, for the most part, the *t*-test results were significant, and KTmaxnorm > zero, suggesting a statistically detectable difference between alveolar and velar attempts. KTmax values were, on average, much larger during intervention, averaging 10.08 mm ( $SD = 4.86$ ). Postintervention, all children produced dorsal gestures for which KTmaxnorm was within the full range of TD children, consistent with our hypothesis, though for two children (03F and 04M) these values were low. In absolute terms, both maxima were only 6 mm. Perhaps, contrary to our expectations, an independent-samples *t* test showed a group difference between the children with SSD postintervention ( $M = 7.1$  mm,  $SD = 2.47$  mm) and

**Figure 7.** Example average tongue splines for /k/ (green) and /t/ (blue) showing a small KTmax difference but a large anterior difference in tongue shape. Anterior is to the right.





the TD children's norms ( $M = 11.9$  mm,  $SD = 3.03$  mm),  $t(33) = 3.368$ ,  $p = .002$ .

For 04M, the value at 6 weeks postintervention fell just outwith the range for TD children, at 5.6 mm, compared to a minimum for TD children of 5.9 mm. For this particular child, this small mean difference was almost functionally sufficient, because he was able to achieve perceptually acceptable productions in 61% of targets.

To summarize, while the KTmaxnorm measure distinguishes well between clearly incorrect and clearly correct productions (as evidenced by the values before and after intervention), values during intervention were variable, suggesting motor instability. We also observed that there was no straightforward relationship between the size of the dorsal differentiation on this measure and the transcription. For example, 15M has similar values for [k] and [q]. It is worth remembering here that the KTmaxnorm measure is a difference measure that quantifies separation of tongue surfaces in the dorsal region, rather than correctness of /k/ articulations, via an analysis of tongue shape. It is not a direct measure of degree and nature of velar contact. However, values outwith the normal range can alert the researcher or clinician to the possibility that the articulation of the velar–alveolar contrast was unusual in some way: KTmaxnorm values (especially when  $t$  tests are significant) may suggest either ULGs or very retracted dorsal articulations. Visualizing the tongue shapes allows a qualitative analysis on an individual basis, as shown in Figures 6 and 7.

## Discussion

While velar fronting is generally understood as a phonological process, there is evidence that some children with SSDs have difficulty with the articulatory gestures required to clearly differentiate alveolar and velar stops. A small number of studies (Cleland et al., 2017; Gibbon, 1999; McAllister Byun, 2012) show that, for children with perhaps more persistent velar fronting, attempts at both alveolars and velars might be realized as ULGs involving increased contact between the tongue tip/blade and dorsum and the hard palate. If this is the case, then persistent velar fronting might be considered an articulatory difficulty with phonological consequences. Comparison articulatory data from TD children is lacking, or confined mainly to EPG studies (e.g., Cheng et al., 2007). However, ultrasound is now being used in more intervention studies due to its relatively low cost and easy application to larger numbers (Cleland, in press). Despite this, prior to this study, few norms for English-speaking children were available.

Moreover, ultrasound is arguably a better tool for imaging dorsal articulations as it is easily able to image post-velar articulations, unlike EPG. This ultrasound study provides norms for the degree of velar and alveolar differentiation of a dorsal gesture during stop production, using data from 30 English-speaking children who had already acquired a typical English /t-/k/ contrast, and who had never received speech therapy. While a degree of individual variability was evident, in general, children in the primary school

years show clear differences in tongue shape for /t/ and /k/ in corner vowel contexts. The maximum radial difference between /k/ and /t/ is larger in /a/ and /o/ vowel contexts than in an /i/ context. This is clearly illustrated in the averaged tongue shapes in Figure 4. We intend this figure to be a useful tongue-shaped template for speech-language pathologists who are working with children who have difficulty acquiring velars, whether this is using ultrasound visual biofeedback or another technique. The diagrams also serve as a reminder for clinicians of the co-articulatory effect vowels have on consonants. We speculate that it may be more difficult for children to differentiate /k/ from /t/ in an /i/ context than in an open or back vowel context at the initial stages of intervention due to the gestures being spatially closer to each other in an /i/ context (Cleland et al., 2015).

This study proposes a measure that compares alveolar and velar gestures directly, and quantifies dorsal raising difference between an active closure constriction achieved by the tongue dorsum and the passive dorsum not forming a phonological constriction for /t/. Developmental English norms were reported for both a maximum radial difference and the area of a crescent limited by anterior and posterior crossovers of /k/ and /t/. In TD English-speaking children, the area and linear difference were found to correlate.

One surprising finding was a lack of correlation between chronological age and area or radial difference. While some growth in the vocal tract would be expected in this age range, it is relatively stable compared to the accelerated growth in early childhood (Vorperian et al., 2005). However, in an EPG study, Cheng et al. (2007) found that younger children had increased tongue–palate contact for /t/ compared to older children and adults. In ultrasound, we could speculate that this would show as a smaller KTmax or area. However, we did not see this pattern. It is worth noting that the data used in this study were from children aged 5;8–12;10, but there was not an even spread of ages ( $M = 9.5$ ;  $SD = 2.04$ ), with only seven children under the age of 8 years and 18 children in the 8–12 age range. It is, therefore, possible that, if data from more younger children had been available, we would have seen a difference. This is clearly a limitation in this study, and it would be useful to have norms for a larger group of children and also for younger children in the preschool years who are in the process of acquiring the contrast. Data from younger children would be a useful comparison for children with SSD to determine whether these children show motor immaturity or more unusual patterns, notwithstanding the difficulties comparing older children to those with potentially much smaller vocal tracts. Nevertheless, the children in this study do represent the typical ages of children undergoing ultrasound visual biofeedback (U-VBF) of over 6 years (Sugden et al., 2019). One small study using U-VBF to treat velar fronting in preschool children found that it was not particularly helpful in that age group (Heng et al., 2016). Longitudinal information from TD children would also be very useful to understand how coronal/dorsal differentiation might change over time and to determine how variable productions are.

In a second study reported here, we measured the maximum radial difference between velars and alveolars

in children with persistent velar fronting, as they were undergoing U-VBF over the course of around 26 weeks. We predicted that, by analyzing the velar gestures as they are acquired, we might gain insight into whether velar fronting might arise from a motoric deficit consistent with the views of both Gibbon (1999) and McAllister Byun (2012). For all these children, before intervention, there was a clear /k/-/t/ merger, reflected in values of KTmax(norm) near to zero and a lack of significant *t* tests for a difference between the tongue shapes for /k/ and /t/. Early in intervention, most of the children became able to achieve an articulation that required them to raise the back of the tongue—although some of the productions involved were not correct, but were transcribed as velar or even uvular fricatives. It is important to remember that we analyzed only successful attempts at dorsal articulations: These results do not include “incorrect” productions, that is, clearly alveolar productions. This is important because children do not make quick categorical shifts toward dorsal productions. Rather, there is evidence of considerable difficulty achieving an appropriate gesture in many of the children, with many attempts fronted to [t] within sessions, especially initially.

In a motor-learning paradigm, learning a new movement is at first slow, inconsistent, and controlled consciously (Weaver, 2015), and this is reflected in the KTmaxnorm values, which vary over time. These results therefore support an articulatory, rather than purely phonological, perspective on remediation. Of course, it might be argued that this effect is amplified here because the children were enrolled in a course of U-VBF, which is a motor-based intervention (Sugden et al., 2019). After all, in this strategy, clinicians might be asking speakers to slow down and focus explicitly on movements, rather than to focus on phonological contrasts. We also see, for some children, articulatory overshoot, with uvular gestures produced instead of velars. It is not clear whether this is a motor control difficulty per se, caused by children exaggerating initial movements, or whether it is due to the knowledge of performance positive reinforcement given to the child when they produce these gestures at the outset of intervention. This overshoot may be in fact due to a primary goal of producing audible contrast or motor patterns that are clearly distinct from previous errors, rather than an inability to produce velar constrictions per se. Over time, however, we see that all of these children began to achieve productions with lower KTmaxnorm values, transcribed as [k], showing that, during intervention, the articulatory gestures became more efficient at achieving velarity and at distinguishing /k/ from /t/, as expected. Despite this, we were surprised to find a significant group difference in KTmax between typical children and children with SSD postintervention, with values being around 40% smaller in the group with SSD. While we are aware of the limitations of the small sample size, it was noted that some of the children produced ULGs for both /t/ and /k/ prior to intervention and continued to produce ULGs for /t/ postintervention. This would account for the smaller KTmax values in this group. It is worth noting that none of the children had completely eliminated the velar fronting pattern at this time

point (Cleland et al., 2019) and we, therefore, speculate that these children continued to have considerable difficulty in the longer term with the gestures required. Perhaps the intervention dosage or intensity was inadequate. Hitchcock et al. (2019) found a significant positive relationship between U-VBF treatment intensity and outcomes, and it is, therefore, possible that the once-per-week treatment these children received was insufficient to completely remediate the disorder. Unfortunately, long-term follow-up was not available for these children.

Our finding of variability in KTMax values during intervention and the evidence of ULGs for alveolar targets, as predicted by both Gibbon (1999) and McAllister Byun (2012), suggests that at least, for children with more persistent velar fronting, this substitution is due to articulatory or motor constraints that are encoded within the child's phonological system. In other words, both phonetic and phonological aspects to the disorder and its remediation can be helpful for our understanding, rather than one or the other exclusively. It is worth noting that typically U-VBF is conceptualized as a motor-based treatment, which involves structured practice of a new articulation (Cleland & Preston, 2021). In this sense, it is similar to traditional articulation intervention where the child practices the sound in error and does not contrast it with other phonemes. However, most studies of U-VBF involve working on distortion-type errors, such as distorted rhotics where there is no merger of phonemes (Sugden et al., 2019). In cases, such as velar fronting, where there is a merger of two or more phonemes, it is likely that a motor-based approach that incorporates a phonological contrast element might be most successful at avoiding overgeneralization. A large body of research suggests that, for children with merger-type errors, phonological intervention leads to greater generalization than articulation approaches alone (e.g., Lousada et al., 2013). It is worth noting, however, that the children in this study had all had prior intervention using phonological approaches that had not been successful. Therefore, for children with more persistent velar fronting, who show evidence of articulatory difficulties, we suggest a motor-based approach such as U-VBF combined with a phonological approach such as minimal pairs intervention (Cleland & Preston, 2021) to both practice the articulatory movements required for producing velars at adequate dosage while encouraging generalization across the child's phonological system.

### Limitations

This study involved secondary analysis of an open access corpus of data (Eshky et al., 2018). While this has advantages in terms of allowing other researchers to access the data and apply the same or different measures, there are clear disadvantages insofar as the data set was not specifically designed to answer our research questions. If new data were to be collected, then we would suggest that a larger sample of TD children, in clearly defined age groups, would be useful. It would also be very helpful to have multiple repetitions of /t/ and /k/ to determine the stability of

gestures within speaker. For the children with SSD, since we took data from real intervention sessions, we did not have consistent word lists during each session: It would have been particularly useful to have multiple attempts at /k/ and /t/ in corner vowel contexts. Nevertheless, the data we analyzed comprise attempts during prepractice as well as practice (Maas et al., 2008), and to our knowledge, this has not been subject to articulatory analysis in any other ultrasound intervention study. This provides a unique insight into the process of learning new articulatory gestures.

Finally, there were some limitations to the measures we designed. First, it was not possible to reliably compute the KT crescent area in the children with SSD as often the tongue shapes were unusual and there were not obvious crossing points, which are needed to define the end points of the crescent. Moreover, the absolute values of KMax could be low even in cases where qualitatively the tongue shapes were very different in the children with SSD. Figure 7 shows a retracted attempt at /k/ where the difference within the crescent shape is low, yet the tongue shapes are very different. It is clear that the main difference between the gestures here is within the anterior region. We, therefore, suggest that a measure that sums the maximum radial difference within the two regions might be useful. While other measures, such as DEI (Zharkova, 2013), can be found in the literature, these have not yet been applied to uvular articulations or ULGs in disordered speech. It is therefore not possible to comment on whether these other measures perform better in atypical speech. Space considerations prevent us from applying DEI to the data used here, and we leave this for future work. It is also important to note that our measures require probe stabilization or correction for probe movement. While this does allow us to take absolute measures between gestures (which might be applied to other techniques such as magnetic resonance imaging or electromagnetic articulography), it does limit jaw movement somewhat and some children may find the headset used here uncomfortable.

## Conclusions

To summarize, this study comprised two parts: In the first part, we present norms for dorsal and coronal differentiation in alveolar and velar stops from 30 TD English-speaking children. We demonstrated that, typically, speakers show a clear difference between /k/ and /t/ productions in corner vowel contexts, but the difference within a high-vowel context is much smaller. Contrary to our expectations, the difference between /k/ and /t/ did not increase with age, although data from a larger group of children stratified for age would be required to confirm this finding. We also provide averaged tongue-shaped diagrams to illustrate these productions, which we hope will provide a useful reference for clinicians and researchers.

In a second study, we explored the usefulness of the linear dorsal differentiation norms for children with persistent velar fronting. Before intervention, all of these children showed a /k/-/t/ merger while, after intervention, the children showed /k/-/t/ differentiation similar to TD children. During

intervention, there was evidence of variable productions at the articulatory level, including uvular articulations and ULGs. This suggests that the substitution of more anterior consonants for velar stops in children with persistent velar fronting is due to articulatory or motor constraints that are encoded within the child's phonological system. We suggest that this is best remediated with a motor-based approach that also incorporates contrast-based elements.

## Acknowledgments

This study was funded by grants from the Engineering and Physical Sciences Research Council (EP/I027696/1) and the Chief Scientist Office of Scotland (ETM-402) awarded to Joanne Cleland and James M. Scobbie. Thank you to the children and their parents who provided the data for the UltraSuite Corpus. Thank you to Zoe Roxburgh for collecting the data from the children with speech sound disorders and to Cornelia Heyde for annotating the typically developing data.

## References

- Articulate Instruments.** (2012). Articulate Assistant Advanced User Guide: Version 2.16. Articulate Instruments Ltd.
- Articulate Instruments Ltd.** (2014). Articulate Assistant Advanced Ultrasound Module User Manual, Revision 2.14. In Articulate Instruments Ltd.
- Baghban, K., Zarifian, T., Adibi, A., Shati, M., & Derakhshandeh, F.** (2020). The quantitative ultrasound study of tongue shape and movement in normal Persian speaking children. *International Journal of Pediatric Otorhinolaryngology*, 134, 110051. <https://doi.org/10.1016/j.ijporl.2020.110051>
- Cheng, H. Y., Murdoch, B. E., Goozee, J. V., & Scott, D.** (2007). Electropalatographic assessment of tongue-to-palate contact patterns and variability in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research*, 50(2), 375–392. [https://doi.org/10.1044/1092-4388\(2007\)027](https://doi.org/10.1044/1092-4388(2007)027)
- Cleland, J.** (in press). Ultrasound tongue imaging. In M. J. Ball (Ed.), *Manual of clinical phonetics*. Routledge.
- Cleland, J., & Preston, J.** (2021). Biofeedback interventions. In A. L. Williams, S. McLeod, & R. McCauley (Eds.), *Interventions for speech sound disorders*. (pp. 573–600) Brookes.
- Cleland, J., Scobbie, J. M., Heyde, C., Roxburgh, Z., & Wrench, A. A.** (2017). Covert contrast and covert errors in persistent velar fronting. *Clinical Linguistics & Phonetics*, 31(1), 35–55. <https://doi.org/10.1080/02699206.2016.1209788>
- Cleland, J., Scobbie, J. M., Roxburgh, Z., Heyde, C., & Wrench, A.** (2019). Enabling new articulatory gestures in children with persistent speech sound disorders using ultrasound visual biofeedback. *Journal of Speech, Language, and Hearing Research*, 62(2), 229–246. [https://doi.org/10.1044/2018\\_JSLHR-S-17-0360](https://doi.org/10.1044/2018_JSLHR-S-17-0360)
- Cleland, J., Scobbie, J. M., & Wrench, A. A.** (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics*, 29(8–10), 575–597. <https://doi.org/10.3109/02699206.2015.1016188>
- Davis, B. L., & MacNeilage, P. F.** (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38(6), 1199–1211. <https://doi.org/10.1044/jshr.3806.1199>
- Dodd, B., Reilly, S., Ttofari Eecen, K., & Morgan, A. T.** (2018). Articulation or phonology? Evidence from longitudinal error data. *Clinical Linguistics & Phonetics*, 32(11), 1027–1041. <https://doi.org/10.1080/02699206.2018.1488994>



- Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic evaluation of articulation and phonology (DEAP)*. The Psychological Corporation.
- Dunn, D., Sewell, D., & J., & Styles, B. (1997). *British Picture Vocabulary Scale 3*. GL Assessment.
- Eshky, A., Ribeiro, M. S., Cleland, J., Richmond, K., Roxburgh, Z., Scobbie, J., & Wrench, A. (2018). UltraSuite: A repository of ultrasound and acoustic data from child speech therapy sessions. *Proceedings of the 19th Annual Conference of the International Speech Communication Association (ISCA)*, 2-6 September 2018, Hyderabad, India, 2018, 1888–1892. <https://doi.org/10.21437/Interspeech.2018-1736>
- Frisch, S. A., & Wodzinski, S. M. (2016). Velar-vowel coarticulation in a virtual target model of stop production. *Journal of Phonetics*, 56, 52–65. <https://doi.org/10.1016/j.wocn.2016.01.001>
- Gibbon, F. E. (1999). Undifferentiated lingual gestures in children with articulation/phonological disorders. *Journal of Speech, Language, and Hearing Research*, 42(2), 382–397. <https://doi.org/10.1044/jslhr.4202.382>
- Gibbon, F. E., & Wood, S. E. (2002). Articulatory drift in the speech of children with articulation and phonological disorders. *Perceptual and Motor Skills*, 95(1), 295–307. <https://doi.org/10.2466/pms.2002.95.1.295>
- Gibbon, F. E., Yuen, I., Lee, A., & Adams, L. (2007). Normal adult speakers' tongue palate contact patterns for alveolar oral and nasal stops. *Advances in Speech Language Pathology*, 9(1), 82–89. <https://doi.org/10.1080/14417040600954824>
- Heng, Q., McCabe, P., Clarke, J., & Preston, J. L. (2016). Using ultrasound visual feedback to remediate velar fronting in preschool children: A pilot study. *Clinical Linguistics & Phonetics*, 30(3–5), 382–397. <https://doi.org/10.3109/02699206.2015.1120345>
- Hitchcock, E. R., Swartz, M. T., & Lopez, M. (2019). Speech sound disorder and visual biofeedback intervention: A preliminary investigation of treatment intensity. *Seminars in Speech and Language*, 40(2), 124–137. <https://doi.org/10.1055/s-0039-1677763>
- Lousada, M., Jesus, L., Capelas, S., & Joffe, V. (2013). Phonological and articulation treatment approaches in Portuguese children with speech and language impairments: A randomized controlled intervention study. *International Journal of Language & Communication Disorders*, 48(2), 172–187. <https://doi.org/10.1111/j.1460-6984.2012.00191.x>
- Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17(3), 277–298. [https://doi.org/10.1044/1058-0360\(2008/025\)](https://doi.org/10.1044/1058-0360(2008/025))
- McAllister Byun, T. (2012). Positional velar fronting: An updated articulatory account. *Journal of Child Language*, 39(5), 1043–1076. <https://doi.org/10.1017/S0305000911000468>
- McAllister Byun, T., Buchwald, A., & Mizoguchi, A. (2016). Covert contrast in velar fronting: An acoustic and ultrasound study. *Clinical Linguistics & Phonetics*, 30(3–5), 249–276. <https://doi.org/10.3109/02699206.2015.1056884>
- McLeod, S., & Baker, E. (2017). *Children's speech: An evidence-based approach to assessment and intervention*. Pearson Education.
- McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology*, 27(4), 1546–1571. [https://doi.org/10.1044/2018\\_AJSLP-17-0100](https://doi.org/10.1044/2018_AJSLP-17-0100)
- Melo, R. M., Mota, H. B., & Berti, L. C. (2017). The contrast between alveolar and velar stops with typical speech data: Acoustic and articulatory analyses. *CoDAS*, 29(3). <https://doi.org/10.1590/2317-1782/20172016117>
- Scobbie, J. M., Gibbon, F., Hardcastle, W., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, 5, 194–207.
- Scobbie, J. M., Lawson, E., & Stuart-Smith, J. (2012). Back to front: A socially-stratified ultrasound tongue imaging study of Scottish English /u/. *Rivista di Linguistica/Italian Journal of Linguistics, Special Issue: Articulatory Techniques for Sociophonetic Research*, 24(1), 103–148.
- Scobbie, J. M., Wrench, A., & van der Linden, M. (2018). Head-probe stabilisation in ultrasound tongue imaging using a head-set to permit natural head movement. In *Proceedings of the 8th International Seminar on Speech Production* (pp. 373–376).
- Stoel-Gammon, C. (1996). On the acquisition of velars in English. In B. Bernhardt, J., Gilbert, & D., Ingram (Eds.), *Proceedings of the UBC International Conference on Phonological Acquisition* (pp. 201–214). Cascadia Press.
- Sugden, E., Lloyd, S., Lam, J., & Cleland, J. (2019). Systematic review of ultrasound visual biofeedback in intervention for speech sound disorders. *International Journal of Language & Communication Disorders*, 54(5), 705–728. <https://doi.org/10.1111/1460-6984.12478>
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S. (2005). Development of vocal tract length during early childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, 117(1), 338–350. <https://doi.org/10.1121/1.1835958>
- Weaver, J. (2015). Motor learning unfolds over different timescales in distinct neural systems. *PLOS Biology*, 13(12), e1002313. <https://doi.org/10.1371/journal.pbio.1002313>
- Wells, J. C. (1982). *Accents of English: Volume 1*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511611759>
- Whalen, D., Iskarous, K., Tiede, M. K., & Ostry, D. J. (2004). HOCUS: The Haskins optically-corrected ultrasound system for measuring speech articulation. *The Journal of the Acoustical Society of America*, 115(5), 2632–2632. <https://doi.org/10.1121/1.4784869>
- Zharkova, N. (2013). Using ultrasound to quantify tongue shape and movement characteristics. *The Cleft Palate–Craniofacial Journal*, 50(1), 76–81. <https://doi.org/10.1597/11-196>
- Zharkova, N. (2019). Differentiating tongue shapes for alveolar-postalveolar and alveolar-velar contrasts. *Speech Communication*, 113, 15–24. <https://doi.org/10.1016/j.specom.2019.08.004>